Polarization-Enhanced Underwater Detection Method for Multiple Material Targets Based on Deep-Learning

Guochen Wang, Jie Gao, Yubin Chen, Xin Wang, Jiangtao Li, Khian-Hooi Chew, and Rui-Pin Chen

***Abstract*—Underwater target detection is an essential topic in the applications of underwater exploration. However, underwater target detection faces serious challenges, such as complex scattering, low visibility, and target clutter. Here a polarization-enhanced underwater multiple material target detection method is proposed to address these challenges. The similarity principle of locally backscattered polarization features is utilized to suppress the influence of backscattered light. Our target detection model combines polarization gradient and edge detection techniques to optimize the detection process, enabling superior target detection and feature extraction. Experimental results indicate that our method has significantly enhanced the detection performance in multiple (overlapping or nonoverlapping) material targets, especially in high turbid underwater scattering environments. This research provides a promising new approach for polarized target detection in underwater environments and opens up new possibilities for underwater multiple-material target detection.**

***Index Terms*—Backscattering, deep learning, polarization, target detection.**

I. Introduction

DUE to complex scattering, low visibility, and cluttered targets, target detection in underwater environments poses significant challenges [1]. Traditional imaging techniques often fail to capture the details and surface characteristics of underwater targets, making it difficult to detect and recognize the material features of targets [2]. Recently, polarization imaging has shown significant advancements in target material detection [3]. Compared to traditional intensity imaging, polarization imaging can provide more information and enhanced contrast, exhibiting better adaptability and robustness in complex scattering environments [4], [5]. In particular, polarization imaging can provide richer surface information and enhanced detection capabilities for material identification by different polarization component images [6], [7]. M. Dubreuil et al. improved detection performance by estimating targets using two orthogonal polarization images [8]. Tang et al. proposed a bilinear polarization discriminator for detecting metal targets [9]. Sun et al. introduced the distribution of polarization information as a feature factor to distinguish metal targets from natural backgrounds [10]. Shi et al. proposed a polarization analysis model with shadow functionality to differentiate the polarization characteristics of metal and non-metal targets [11]. Usmani et al. have made good progress in both image restoration and target detection in low-light environments using polarimetric three-dimensional integral imaging techniques [12], [13]. However, most works about polarization-based target detection methods mainly focus on distinguishing targets in non-scattering environments. Signal detection [14], [15] and three-dimensional imaging [16], [17] in scattering environments have been reported. Nevertheless, detecting and recognizing multiple targets with various categories of material in complex scattering environments (such as turbid underwater) is still an important topic in target exploration due to its fundamental physics interest and potential applications. It requires more generalization and anti-interference capabilities to suppress backscattering influence and detect multiple material targets in polarization imaging-based target detection models.

In this work, we propose a method to enhance the polarization features of target surfaces and edges by suppressing the backscattering influence for detecting multiple materials of targets in turbid underwater environments. The impact of backscattering light is suppressed by comparing the degree of linear polarization (DOLP) and the angle of polarization (AOP) information of adjacent pixels, according to the local similarity principle of backscattering polarization information in polarized images [18]. In particular, the comparison between the polarization information (AOP and DOLP) of adjacent pixels can effectively reduce the influence of the scattering effect of particulate matter in a turbidity environment is adopted, for the first time to our best knowledge, for the target detection in a turbid underwater environment. According to the physical detection model of polarization characteristics by subtracting the AOP and DOLP between neighboring pixels, the neural network is

Manuscript received 15 August 2023; revised 4 October 2023; accepted 17 October 2023. Date of publication 23 October 2023; date of current version 30 October 2023. This work was supported in part by Zhejiang Provincial Key Research and Development Program under Grant 2022C04007 and in part by the National Natural Science Foundation of China under Grant 11874323. *(Corresponding author: Rui-Pin Chen.)*

Guochen Wang, Jie Gao, Yubin Chen, Xin Wang, Jiangtao Li, and Rui-Pin Chen are with the Key Laboratory of Optical Field Manipulation of Zhejiang Province, Department of Physics, Zhejiang Sci-Tech University, Hangzhou 310018, China (e-mail: sdkj_wgc@163.com; gaojie657@163.com; 202230106280@mails.zstu.edu.cn; 202230106320@mails.zstu.edu.cn; 849720672@qq.com; chenrp@zstu.edu.cn).

Khian-Hooi Chew is with the Key Laboratory of Optical Field Manipulation of Zhejiang Province, Department of Physics, Zhejiang Sci-Tech University, Hangzhou 310018, China, and also with the Zhejiang Expo New Materials Company Ltd., Wenzhou 325802, China (e-mail: khchew2@hotmail.com).

Digital Object Identifier 10.1109/JPHOT.2023.3326158

constructed with the input of 4 polarization component images to map physical characteristic parameters of the target. Polarization gradient and edge detection techniques are employed to effectively trace the boundaries and contours of underwater targets in the target detecting network. Experimental results exhibit the superiority of our method in the detection accuracy of multiple material targets compared to state-of-the-art methods, especially in high-turbidity underwater scattering environments. The physical model demonstrated in this approach in this manuscript is effective for both linear and circular polarizations. We focus on the linear polarizations for the target material detection in scattering environments in this work because the linear polarizations are common polarization components and are more convenient to measure in practice. This work can be extended to the circularly polarized target material detection with the circularly polarized imaging, provided that the input beam with circular polarization and measurement of polarization information distribution of circularly polarized images. The procedure is similar to that of linear polarization.

II. Principle of Polarization-Enhanced

A. Backscattering Suppression and Edge Enhancement

According to the underwater imaging model, underwater intensity images comprise target signals and backscattered information. The total intensity of the image can be expressed as [18]:

$$I(x,y) = T(x,y) + B(x,y), \tag{1}$$

where $T(x, y)$ represents the target information, and $B(x, y)$ represents the backscattered information. Here we extend the polarization imaging model with Stokes parameters [3]. According to the definition of Stokes parameters, S_0 denotes the total intensity, S_1 represents the intensity of the linearly polarized components in the horizontal and vertical directions, and S_2 represents the intensity of the linearly polarized components in the 45° and 135° directions. Therefore, The S_1 of captured images in CCD can be expressed as the sum of the target reflection light S_1^T and the backscattering light S_1^B. The values of S_2 are the sum of the target reflection light S_2^T and the backscattering light S_2^B, where S_2^T represents the portion of S_2 corresponding to target-reflected light, and S_2^B represents the portion of S_2 corresponding to backscattering light [3]:

$$S_1(x,y) = S_1^T(x,y) + S_1^B(x,y), \tag{2}$$

$$S_2(x,y) = S_2^T(x,y) + S_2^B(x,y). \tag{3}$$

The corresponding AOP and the DOLP can be defined as follows [19]:

$$AOP = \frac{1}{2}\arctan\left(\frac{S_2}{S_1}\right), \tag{4}$$

$$DOLP = \sqrt{S_1^2 + S_2^2}/S_0. \tag{5}$$

The difference between AOP and DOLP between neighboring pixels can be expressed as:

$$d_{AOP} = \frac{1}{2}\arctan\left[\frac{\frac{S_2^T+S_2^B}{S_1^T+S_1^B} - \frac{S_2^{T'}+S_2^{B'}}{S_1^{T'}+S_1^{B'}}}{1+\frac{(S_2^T+S_2^B)(S_2^{T'}+S_2^{B'})}{(S_1^T+S_1^B)(S_1^{T'}+S_1^{B'})}}\right], \tag{6}$$

$$d_{DOLP} = \sqrt{(S_1^T+S_1^B)^2+(S_2^T+S_2^B)^2}/S_0 - \sqrt{(S_1^{T'}+S_1^{B'})^2+(S_2^{T'}+S_2^{B'})^2}/S_0. \tag{7}$$

In turbid underwater environments, backscattering occurs due to the interaction between light and suspended particles [20]. In a homogeneous medium, the distribution of these particles is relatively uniform, resulting in similar scattering effects on adjacent pixels in the image. Therefore, the corresponding backscattering Stokes values of neighboring pixels can be considered approximately equal (similarity principle of locally backscattered polarization) [18], i.e., $S_{0,1,2}^B \approx S_{0,1,2}^{B'}$. In particular, in a highly turbid underwater environment, the backscattered components $S_{0,1,2}^B$ are much larger than that of the reflected component from the target $S_{0,1,2}^T$, i.e., $S_{0,1,2}^B >> S_{0,1,2}^T$ [21], [22]. Under these conditions, (6) and (7) can be approximately represented as:

$$d_{AOP} = \frac{1}{2}\arctan\left[\frac{S_2^T - S_2^{T'}}{S_1^B + \frac{S_1^B S_2 S'_2}{S_1 S'_1}}\right], \tag{8}$$

$$d_{DOLP} = \frac{(S_1^T - S_1^{T'})(S_1 + S'_1) + (S_2^T - S_2^{T'})(S_2 + S'_2)}{S_0(\sqrt{S_1^2+S_2^2} + \sqrt{S_1^2+S_2^2})}. \tag{9}$$

As shown in (8) and (9), it is possible to effectively reduce the impact of the backscattering light in a highly turbid underwater environment by subtracting the Stokes parameters between neighboring pixels. In particular, the higher the turbidity of the scattering medium, the more significant the suppression effect of the backscattering light influence, leading to a better restoration of the reflected light from the target. It is worth noting that the $S_{0,1,2}^T$ undergoes abrupt changes at the boundary between the background and the target, leading the components d_{AOP} and d_{DOLP} to exhibit significant differences at the boundary and highlight the contours of the target. Thus, the detection effect of a target can be improved. Here, an operator Os is constructed in the network to calculate this process with polarimetric images:

$$Os = \sqrt{\frac{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} d_{i,j}^2}{\max(d) - \min(d)}}. \tag{10}$$

A sliding window with a size of $n_1 \times n_2$ and a stride of 1 is created to perform an accumulation calculation traversing the entire image, and a constraint on the polarization scattering difference $d(d_{AOP}, d_{DOLP})$ is imposed within the window.

In addition, feature confusion and boundary blur occur during target detection, especially for overlapping objects, significantly affecting the effectiveness of target detection.

Here the Sobel operator [23] and the Canny edge detection algorithm [24] are introduced, for the first time, to our knowledge,

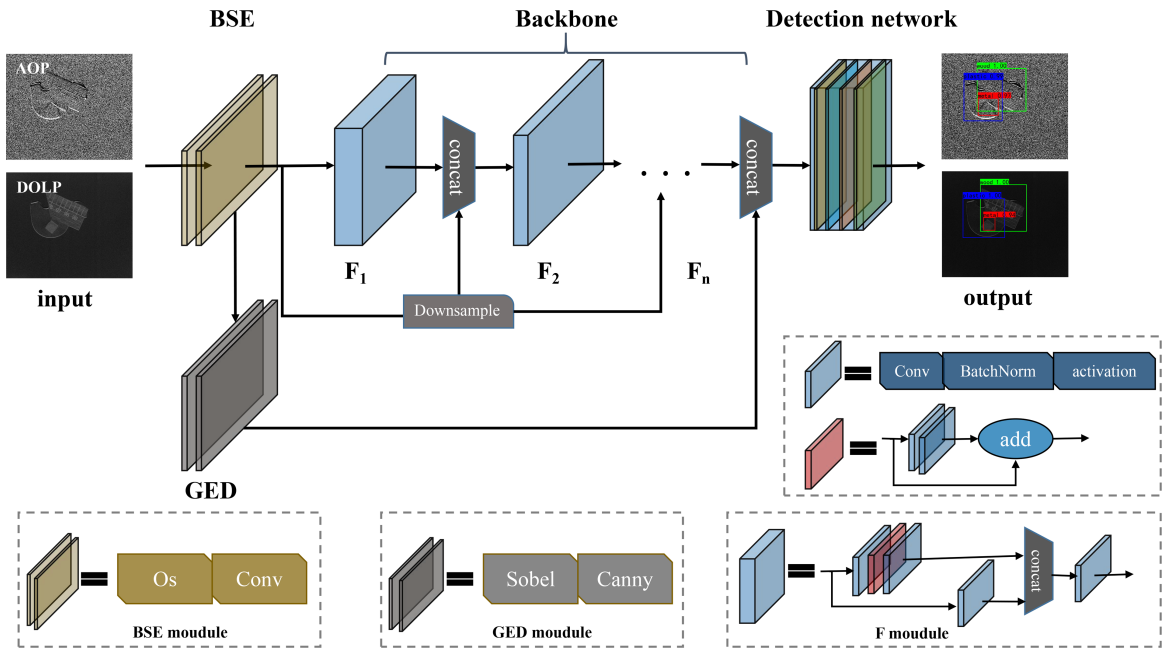


Fig. 1. Network model architecture.

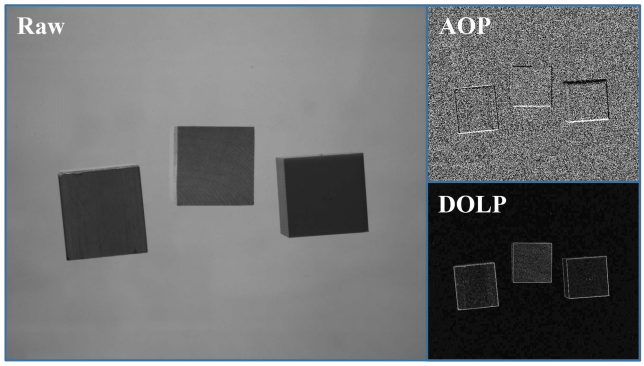
for the gradient and edge detection algorithms with polarimetric images to address the issues. The gradient magnitude and direction for each pixel are calculated using the Sobel operator, whereas the Canny edge detection algorithm obtains the gradient magnitude and direction variation.

B. Polarization-Enhanced Target Detection Network

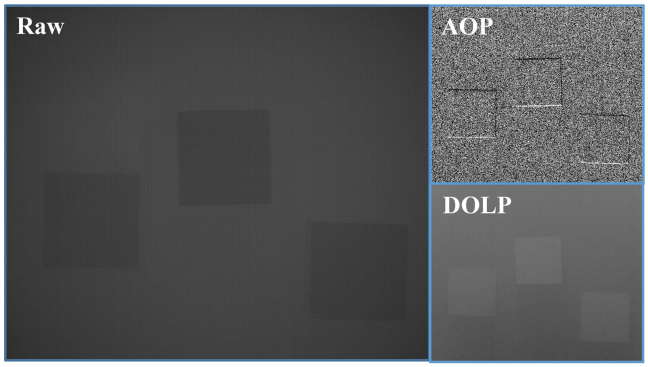
The entire model network architecture is shown in Fig. 1. In this work, the designed BSE and GED modules are components of an object detection model suitable for polarized images and can be applied to the majority of object detection models. The backbone consists of F modules. F modules represent convolutional blocks used to extract features in the network. Typically, F modules consist of multiple convolutional layers (or along with other network structures). Herein three mainstream object detection frameworks (EfficientDet, Faster R-CNN, and YOLOv4) combining the PFE modules are employed to conduct the target detection. The specific number of F modules and the input size of images depend on the chosen object detection frameworks. The backbone network of EfficientDet consists of 16 MBConv blocks with an input size of 1536×1536 and a learning rate of 0.256 and 100 iterations [25]. The backbone network of Faster R-CNN is VGG16, consisting of 13 Conv layers and 3 FC layers, with an input size of 608×608 and a learning rate of 0.001 and 100 iterations [26]. The backbone network of YOLOv4 consists of 23 CSP blocks, with an input size of 608×608 and a learning rate of 0.001 and 100 iterations [27]. Specifically, the BSE module and the GED module are added to the backbone network of our proposed network to enhance the feature extraction capability of the backbone network. The polarization feature enhancement (PFE) module includes Backscatter Suppression Enhancement (BSE) and Gradient and Edge Detection (GED) modules. Firstly, the processing for the suppression of backscattering and the enhancement of the target reflection signal is carried out using the Os operator in the BSE module with the input of AOP and DOLP images. The output of the BSE module is simultaneously passed to the feature extraction layer F_1 and the GED module for processing. Next, in the GED module, the Sobel operator and the Canny edge detection algorithm operator are employed to enhance the edge contours of the targets. The processed images from the GED module are directly output to the F_n layer without passing through the feature extraction layers due to the possible loss of target content information. The feature map generated by the feature extraction layer of the backbone network F_1 is concatenated with the BSE layer and passed to the next feature extraction layer F_2. During this process, a downsampling method is used to maintain the consistent data dimensions. This process is repeated until the last feature processing layer F_n, and the final output feature map is concatenated with the edge feature map generated by the GED module. After the model is trained, the feature map of the input image can be achieved by the network to detect multiple material targets in a turbid underwater environment.

III. Experimental Setup

In the experiments, skim milk was used as the scattering medium to simulate turbid underwater environments. The turbidity of the underwater environment ranged from 0 NTU to 80 NTU. The underwater polarization information of the targets was captured using a Division of Focal Plane (DoFP) camera (LUCID, PHX050S-P, 2448×2048 resolution). This camera can simultaneously capture images with different polarization components (0°, 45°, 90°, 135° linear polarization), from which AOP



(a)



(b)

Fig. 2. Underwater polarimetric images and corresponding AOP and DOLP images: (a) Clear underwater, (b) turbid underwater. (From left to right in each plot: Plastic, wood, metal).

and DOLP images are computed, as shown in Fig. 2. The camera captures multiple material targets with various positions and numbers in underwater environments with different turbidities. The material categories of targets in this work included metal, plastic, and wood. A dataset of 12000 underwater polarization images was obtained, and randomly divided into a training set and a validation set in the ratio of 9:1. The training of the model was conducted on an NVIDIA RTX 3090ti GPU.

The images extracted from the network after being processed by the BSE module and GED module are compared with corresponding AOP and DOLP images to visually demonstrate the effectiveness of BSE and GED modules in suppressing backscattering impact and enhancing edge features, as shown in Fig. 3. The image feature qualities are greatly improved, and the BSE module effectively suppresses the effect of backscattering light. The edge features are further processed in the GED module, highlighting the contours of overlapping targets, as shown in Fig. 3. Here AOP-BSE and DOLP-BSE images were obtained by Os operation, AOP-GED and DOLP-GDE images were obtained by Sobel and Canny operation.

IV. Results and Discussion

We trained object detection models, including EfficientDet [25], Faster R-CNN [26], and YOLOv4 [27] with the underwater

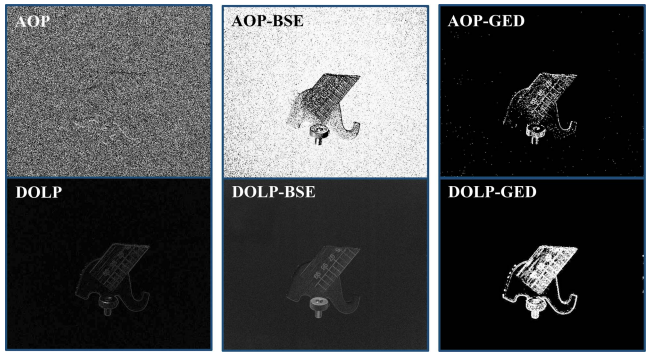


Fig. 3. PFE effects with AOP and DOLP. The left column: AOP and DOLP images; the middle column: The images after passing the BSE module; the right column: The images after passing the GED module.

polarization dataset and demonstrated their detection performance in turbid underwater environments. The selected object detection models (EfficientDet, Faster R-CNN, YOLOv4) were initialized using pre-trained weights. The dataset is annotated by marking the boundaries of the targets and labeling their respective categories. Data augmentation techniques, including rotation, cropping, and flipping, were applied to increase the diversity of the training set and enhance the model's robustness. During training, the features of the loaded data were enhanced by the invoked BSE and GED modules, and then propagated through the subsequent layers of the network. The performances of each model are iteratively optimized by using its respective loss function. Upon completion of training, weight files were finally obtained to validate the model's effectiveness. Fig. 4(a) demonstrates the detection results of each model without a PFE module for the multiple material targets in a turbid underwater environment (NTU = 50). The well-known Faster R-CNN model, renowned for its detection accuracy, fails to accurately detect the material information of the targets in turbid underwater environments. The lightweight EfficientDet model and real-time YOLOv4 model perform even worse in detecting multiple material targets. Fig. 4(b) illustrates the detection results of the corresponding object detection models, including the PFE module for the multiple material targets in similar turbid underwater environments. Compared to the models without PFE modules, these object detection models with PFE modules show significant improvements in detection performance with significantly increased confidence [28], including accurately locating the positions and recognizing the material of the target objects in turbid underwater environments.

In particular, the overlapped multiple material targets are further tested to demonstrate the generalization and robustness of the proposed method. The target occlusion caused by overlapping remains challenging in almost all object detection tasks [29], especially in turbid underwater environments. The interference of feature confusion and boundary-blurring caused by overlapping targets is a stubborn issue for traditional object detection models, leading to incorrect detection, as shown in Fig. 5(a). However, this issue can be well addressed with our proposed method by refining the edge detection of the target objects

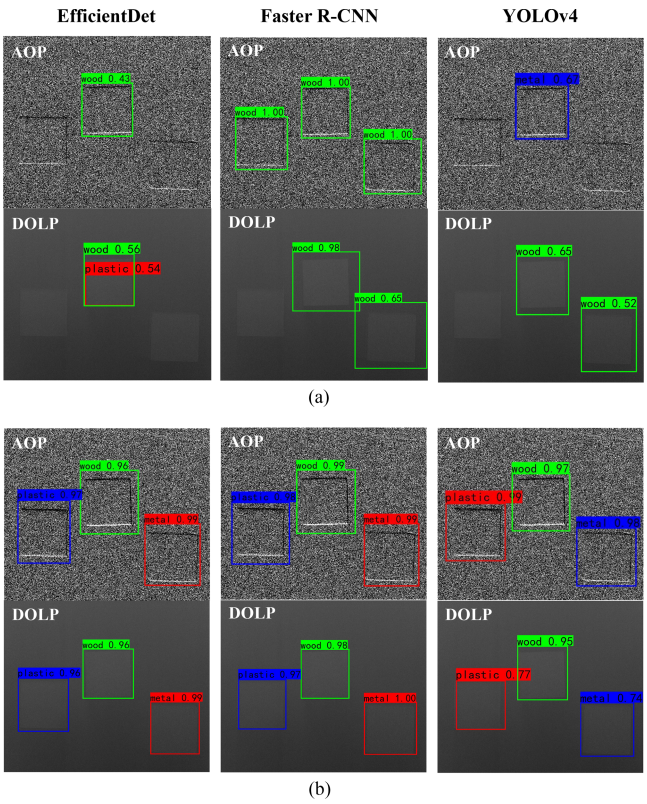


Fig. 4. Comparison of detecting results of traditional target detection models without and with PFE module for multiple material targets in a turbid underwater environment (NTU = 50): (a) Without PFE module; (b) with PFE module. (The labeled information in the figure are the material categories of the target and the confidence level values).

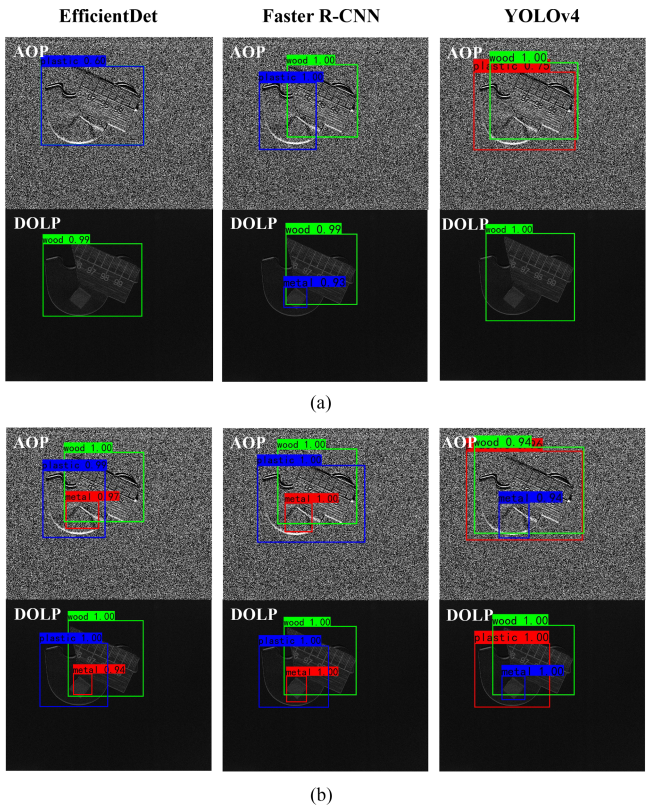


Fig. 5. Comparison of detecting results of traditional target detection models without and with PFE module for overlapped multiple material targets in turbid underwater environment (NTU = 20): (a) Without PFE module; (b) with PFE module. (The labeled information in the figure are the material categories of the target and the confidence level values).

and further adjusting the confidence score and non-maximum suppression (NMS) [30] threshold, as shown in Fig. 5(b). The detection models with PFE modules can achieve excellent performance in detecting overlapping targets in a turbid underwater environment.

The quantitative comparison is further evaluated for the performance of the aforementioned models without and with the PFE module using metrics such as mAP(mean Average Precision), F-Measure, Recall, and Precision [31]. The results are presented in Table I, which shows the values of these metrics for the EfficientDet, Faster R-CNN, and YOLOv4 models without and with PFE modules. The data demonstrates that our proposed PFE method effectively improves the detection performance of the object detection models for multiple material targets in a turbid underwater environment, as evidenced by the higher values of these metrics than those of the original models.

The ablation experiments are conducted to further evaluate the effectiveness of the proposed method. The mAP in Table II represents the average value of different network architectures (EfficientDet, Faster R-CNN, and YOLOv4). The average mAP values of each model (with different combinations of BSE and GED modules) is computed with different network architecture (EfficientDet, Faster R-CNN, and YOLOv4), as the evaluation metric in the ablation experiments. The average performance

TABLE I
PERFORMANCE COMPARISON OF TARGET DETECTION MODELS WITH OR WITHOUT PFE MODULE

Method Criterion	mAP	F-Measure	Recall	Precision
EfficientDet	91.74%	0.88	90.83%	86.01%
EfficientDet-PFE	**99.52%**	**0.98**	**99.24%**	**97.76%**
Faster R-CNN	94.13%	0.89	94.28%	90.57%
Faster R-CNN-PFE	**99.78%**	**0.99**	**99.46%**	**99.62%**
YOLOv4	90.73%	0.86	91.20%	84.73%
YOLOv4-PFE	**99.44%**	**0.98**	**99.24%**	**97.39%**

The bold values/entities indicates these values are obtained by using our methods.

TABLE II
MODEL PERFORMANCE WITH DIFFERENT NETWORK MODULES

Backbone	BSE	GED	mAP
√			91.07%
√	√		97.74%
√		√	93.15%
√	√	√	99.58%

of the network models with different combinations of BSE and GED are presented in Table II. With the BSE or GED module, the network's mAP values increase to 97.74% and 93.15%, respectively, from 91.07%. The network's mAP value can achieve 99.58% with both BSE and GED modules. In addition, regarding the applicability of the model in cluttered scenes of real underwater imaging scenarios, the model can be trained with target images in different situations with different levels of clutter to enhance the generalization of the model's detection effect, achieving real underwater imaging scenarios.

V. Conclusion

This paper proposes a PFE underwater polarization target detection method, combining deep learning techniques to detect multiple material targets in turbid underwater environments. Based on the interference mechanism of backscattering light in polarization parameters, the impact of backscattering light is effectively suppressed by exploiting the local polarization similarity principle of backscattering light, consequently enhancing detection performance. The networks are constructed according to the physical model to detect multiple targets with various categories of material in turbid underwater environments. The polarization gradient and edge detection modules are designed to guide the network to pay more attention to the edge information of the targets in the polarimetric images. Experimental results demonstrate that our proposed method has significant performance for detecting multiple material targets in turbid underwater environments compared to state-of-the-art techniques.

References

[1] Y. Xiang et al., "Underwater polarization imaging recovery based on polarimetric residual dense network," *IEEE Photon. J.*, vol. 14, no. 6, Dec. 2022, Art. no. 7860206.

[2] S. Xu, M. Zhang, W. Song, H. Mei, Q. He, and A. Liotta, "A systematic review and analysis of deep learning-based underwater object detection," *Neurocomputing*, vol. 527, pp. 204–232, Mar. 2023.

[3] Y. Wei, P. Han, F. Liu, and X. Shao, "Enhancement of underwater vision by fully exploiting the polarization information from the Stokes vector," *Opt. Exp.*, vol. 29, no. 14, pp. 22275–22287, Jun. 2021.

[4] Y. Zhao, W. He, H. Ren, Y. Li, and Y. Fu, "Polarization descattering imaging through turbid water without prior knowledge," *Opt. Laser Eng.*, vol. 148, Jan. 2022, Art. no. 106777.

[5] F. Liu et al., "Deeply seeing through highly turbid water by active polarization imaging," *Opt. Lett.*, vol. 43, no. 20, pp. 4903–4906, Oct. 2018.

[6] K. Liu and Y. Liang, "Enhancement of underwater optical images based on background light estimation and improved adaptive transmission fusion," *Opt. Exp.*, vol. 29, no. 18, pp. 28307–28328, Aug. 2021.

[7] E. J. Nunes-Pereira, H. Peixoto, J. Teixeira, and J. Santos, "Polarization-coded material classification in automotive LIDAR aiming at safer autonomous driving implementations," *Appl. Opt.*, vol. 59, no. 8, pp. 2530–2540, Mar. 2020.

[8] M. Dubreuil, P. Delrot, I. Leonard, A. Alfalou, C. Brosseau, and A. Dogariu, "Exploring underwater target detection by imaging polarimetry and correlation techniques," *Appl. Opt.*, vol. 52, no. 5, pp. 997–1005, Feb. 2013.

[9] F. Tang, L. Gui, J. Liu, K. Chen, L. Lang, and Y. Cheng, "Metal target detection method using passive millimeter-wave polarimetric imagery," *Opt. Exp.*, vol. 28, no. 9, pp. 13336–13351, Apr. 2020.

[10] R. Sun, X. Sun, F. Chen, H. Pan, and Q. Song, "An artificial target detection method combining a polarimetric feature extractor with deep convolutional neural networks," *Int. J. Remote Sens.*, vol. 41, no. 13, pp. 4995–5009, Feb. 2020.

[11] H. Shi, Y. Liu, C. He, C. Wang, Y. Li, and Y. Zhang, "Analysis of infrared polarization properties of targets with rough surfaces," *Opt. Laser Tech.*, vol. 151, Jul. 2022, Art. no. 108069.

[12] K. Usmani, G. Krishnan, T. O'Connor, and B. Javidi, "Deep learning polarimetric three-dimensional integral imaging object recognition in adverse environmental conditions," *Opt. Exp.*, vol. 29, no. 8, pp. 12215–12228, Apr. 2021.

[13] K. Usmani, T. O'Connor, and B. Javidi, "Three-dimensional polarimetric image restoration in low light with deep residual learning and integral imaging," *Opt. Exp.*, vol. 29, no. 18, pp. 29505–29517, Aug. 2021.

[14] R. Joshi, G. Krishnan, T. O'Connor, and B. Javidi, "Signal detection in turbid water using temporally encoded polarimetric integral imaging," *Opt. Exp.*, vol. 28, no. 24, Nov. 2020, Art. no. 36033.

[15] Y. Huang, G. Krishnan, T. O'Connor, R. Joshi, and B. Javidi, "End-to-end integrated pipeline for underwater optical signal detection using 1D integral imaging capture with a convolutional neural network," *Opt. Exp.*, vol. 31, no. 2, pp. 1367–1385, Jan. 2023.

[16] M. Cho and B. Javidi, "Three-dimensional visualization of objects in turbid water using integral imaging," *J. Display Technol.*, vol. 6, no. 10, pp. 544–547, Oct. 2010.

[17] M. Cho and B. Javidi, "Peplography—A passive 3D photon counting imaging through scattering media," *Opt. Lett.*, vol. 41, no. 22, pp. 5401–5404, Nov. 2016.

[18] T. Treibitz and Y. Y. Schechner, "Active polarization descattering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 3, pp. 385–399, Mar. 2009.

[19] J. Liang, L. Ren, E. Qu, B. Hu, and Y. Wang, "Method for enhancing visibility of hazy images based on polarimetric imaging," *Photon. Res.*, vol. 2, no. 1, pp. 38–44, Feb. 2014.

[20] Q. Ren, Y. Xiang, G. Wang, J. Gao, Y. Wu, and R.-P. Chen, "The underwater polarization dehazing imaging with a lightweight convolutional neural network," *Optik*, vol. 251, Feb. 2022, Art. no. 168381.

[21] J. Zhou, T. Yang, W. Chu, and W. Zhang, "Underwater image restoration via backscatter pixel prior and color compensation," *Eng. Appl. Artif. Intell.*, vol. 111, May 2022, Art. no. 104785.

[22] B. Huang, T. Liu, H. Hu, J. Han, and M. Yu, "Underwater image recovery considering polarization effects of objects," *Opt. Exp.*, vol. 24, no. 9, pp. 9826–9838, Apr. 2016.

[23] R. Zhang, Q. Xiao, Y. Du, and X. Zuo, "DSPI filtering evaluation method based on Sobel operator and image entropy," *IEEE Photon. J.*, vol. 13, no. 6, Dec. 2021, Art. no. 7800110.

[24] H. Fu, B. Wu, Y. Shao, and H. Zhang, "Perception oriented haze image definition restoration by basing on physical optics model," *IEEE Photon. J.*, vol. 10, no. 3, Jun. 2018, Art. no. 3900816.

[25] M. Tan, R. Pang, and Q. V. Le, "EfficientDet: Scalable and efficient object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 10781–10790.

[26] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[27] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," 2020, *arXiv:2004.10934*.

[28] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 779–788.

[29] Y. Wei, R. Tao, Z. Wu, Y. Ma, L. Zhang, and X. Liu, "Occluded prohibited items detection: An X-ray security inspection benchmark and deocclusion attention module," in *Proc. 28th Assoc. Comput. Machinery Int. Conf. Multimedia*, 2020, pp. 138–146.

[30] Y. He, C. Zhu, J. Wang, M. Savvides, and X. Zhang, "Bounding box regression with uncertainty for accurate object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 2888–2897.

[31] A. M. Roy, R. Bose, and J. Bhaduri, "A fast accurate fine-grain object detection model based on YOLOv4 deep neural network," *Neural Comput. Appl.*, vol. 34, no. 5, pp. 3895–3921, Jan. 2022.